Received 6 June 2000

CLAIMS

1. A method of protein engineering including the steps of:-

(i) creating a computer database which includes a plurality of entries, each said entry corresponding to a description of a location and orientation in 3D space of side chains of amino acid residues of a framework protein, wherein the location and orientation of each side chain is simplified as a Cα-Cβ vector;

(ii) creating a query corresponding to a description of a location and orientation in 3D space of respective side chains of two or more amino acid residues of a sample protein which are required for a function of said sample protein, wherein the location and orientation of each side chain is simplified as a Cα-Cβ vector; and

(iii) searching said database with said query to thereby identify one or more hits wherein at least one of said hits corresponds to a respective said framework protein which has structural similarity to said sample protein.

2. A method of protein engineering including the steps of:-

(i) creating a computer database which includes a plurality of entries, each said entry corresponding to a description of a location and orientation in 3D space of amino acid residues of a framework protein capable of internal disulfide bond formation;

(ii) creating a query corresponding to a description of a location and orientation in 3D space of two or more amino acid residues of a sample protein which are required for a function of said sample protein; and

(iii) searching said database with said query to thereby identify one or more hits wherein at least one of said

hits corresponds to a respective said framework protein which has structural similarity to said sample protein.

3. The method of Claim 1, wherein the framework protein is capable of internal disulfide bond formation.

4. The method of Claim 3, wherein said framework protein is a small cysteine rich protein which comprises 70 amino acids or less, having 2-11 disulfide bonds.

5. The method of Claim 2, wherein said framework protein is a small cysteine rich protein which comprises 70 amino acids or less, having 2-11 disulfide bonds.

6. The method of 5, wherein the location and orientation of a side-chain of each said amino acid residue of said framework protein and the location and orientation of a side-chain of each of said two or more amino acid residues of said sample protein is simplified as a respective Cα-Cβ vector.

7. The method of any one of Claims 1, 3, 4 or 6, wherein the Cα-Cβ vector is in the form of a distance matrix representation.

8. The method of Claim 1 or Claim 2, further including the step of modifying an amino acid sequence of said framework protein which corresponds to a hit, by substituting at least one amino acid residue thereof with at least one amino acid residue of said sample protein to thereby create a modified framework protein.

9. The method of Claim 8, wherein the at least one amino acid residue of said sample protein represents at least a portion of at least one functional region of said sample protein.

10. The method of Claim 9, wherein at least two of the amino acid residues of said sample protein which substitute amino acid residues of said framework protein are non-contiguous in primary sequence.

11. The method of any one of Claims 8-10, wherein the modified framework protein has greater stability than said sample protein.

12. The method of any one of Claims 8-11, wherein the modified framework protein has increased structural similarity to said sample protein.

13. The method of Claim 12, wherein the modified framework protein is capable of exhibiting a function which is either similar to, or inhibitory of, a function of said sample protein.

14. The method of any preceding claim, wherein the sample protein is a cytokine.

15. The method of Claim 14, wherein the cytokine is selected from the group consisting of GH, IL-4, IL-6 and G-CSF.

16. The method of Claim 1 or Claim 2, wherein at step (iii) the hits are ranked according to structural similarity with said sample protein.

17. The method of Claim 1 or Claim 2, wherein searching at step (iii) includes:

(a) identification of said hits by clique detection;

(b) filtering of said hits identified at step (a).

18. A modified framework protein produced according to the method of any one of Claims 9-15.

19. The modified framework protein of Claim 18, which protein is a cytokine mimetic.

20. An engineered protein comprising 70 amino acid residues or less of a framework protein and 2-11 disulfide bonds of said framework protein, together with at least two amino acid residues of another protein which are non-contiguous in primary sequence and represent at least a portion of a functional region of said another protein.

21. The engineered protein of Claim 20, which protein has greater stability than said another protein.

22. The engineered protein of Claim 21, which protein exhibits a function either similar to, or inhibitory of, said another protein.

23. The engineered protein of any one of Claims 20-22, wherein said another protein is a cytokine.

Received 6 June 2000

24. The engineered protein of Claim 23, wherein the cytokine is selected from the group consisting of GH, IL-4, IL-6 and G-CSF.

25. The engineered protein of Claim 24, said engineered protein having an amino acid sequence selected from the group consisting of SCY01, SCY02, SCY03, ERP01, ERP02, ERP03 and VIBO1.

26. The engineered protein of Claim 25, which protein is a cytokine mimetic.

27. A computer program for searching a protein database which comprises a plurality of entries, each said entry corresponding to a distance matrix representation of two or more Cα-Cβ vectors, said program including the steps of:

(i) comparing a query with each said database entry, said query corresponding to a distance matrix representation of two or more Cα-Cβ vectors; and

(ii) identifying hits by clique detection, wherein a hit is defined according to a minumum number of Cα-Cβ vector matches between said query and each said entry.

28. A computer program which filters said hits identified at step (ii) of Claim 27.

29. A computer program according to Claim 27, which program is a VECTRIX program as described herein.

30. A computer program according to Claim 28, which program is a POSTVEC program as described herein.

add B1